

# INCORPORANDO O ATRASO ESCOLAR E AS CARACTERÍSTICAS SÓCIO-DEMOGRÁFICAS NAS TAXAS DE TRANSIÇÃO EDUCACIONAL: UM MODELO DE FLUXO ESCOLAR

Elaine Toldo Pazello<sup>1</sup>  
Reynaldo Fernandes<sup>2</sup>  
Fabiana de Felício<sup>3</sup>

## RESUMO:

O objetivo do artigo é apresentar um modelo de fluxo escolar que a partir de novas formas de estimar as probabilidades de ingresso, aprovação e evasão escolar, contribua com melhores estimativas para a evolução das matrículas no Ensino Fundamental e Médio. A partir de dois anos de PNAD (2002 e 2003) identificam-se as taxas de transição escolar e com base nessas taxas, projeta-se a PNAD até 2009, no que se refere às características educacionais. As principais contribuições do artigo estão no fato das taxas de transição serem condicionais ao atraso escolar, diferentemente dos modelos anteriores que supunham um processo markoviano, o que não parece razoável devido à correlação positiva entre o atraso escolar e a probabilidade de repetência. Além disso, características sócio-demográficas também são consideradas refinando, assim, as previsões de fluxo de escolar conforme mudanças na composição das coortes. De acordo com as estimativas obtidas, as taxas de ingresso e aprovação diminuem com o atraso escolar e aumentam com a escolaridade dos pais; para a evasão, o resultado é o inverso. Os resultados das projeções indicaram o que parece ser um ajuste das matrículas nas séries iniciais ao processo de universalização do acesso a escola que ocorreu ao longo da década passada. Para o ensino médio, as projeções indicaram um crescimento significativo das matrículas que pode ter origem em coortes maiores ou em mudanças de características sócio-demográficas.

**Palavras-Chaves:** fluxo escolar, ingresso, reprovação, evasão, matrículas

**Área da ANPEC:** Área 11 - Economia Social e Demografia Econômica

## ABSTRACT:

The aim of this paper is to introduce a model of student flows. Starting with two years of the Brazilian Household Survey PNAD (2002 and 2003), school transition rates are identified for primary school and high school. These rates are used to project the education-related data in PNAD until 2009. The main contribution of the paper is to condition transition rates on schooling delay, unlike previous studies, which relied on a Markovian model. Furthermore, demographics are also taken into account. The estimates suggest that first grade entrance and approved rates decrease with schooling delay and increase with more educated parents. The projections indicate a smaller enrollment in the early grades of primary school. This may be the consequence of the primary school universalization, which has taken place in the last decade. The projections also indicate a significant rise in enrollment in high school.

**Keywords:** student flows, first grade entrance, retention, dropout, enrollment

Code JEL: I21

---

<sup>1</sup> Do Departamento de Economia da FEA-RP / USP.

<sup>2</sup> Da ESAF e do Departamento de Economia da FEA-RP / USP.

<sup>3</sup> Da ESAF.

# Incorporando o atraso escolar e as características sócio-demográficas nas taxas de transição educacional: Um modelo de Fluxo Escolar

## Introdução

A discussão sobre o fluxo escolar das crianças brasileiras ganhou importância nos anos 90. Até os anos 80, atribuía-se à evasão escolar a responsabilidade pelo fraco acesso ao ensino de 1º grau no Brasil. Este diagnóstico passou a ser questionado no final dos anos 80<sup>4</sup>. O principal ponto levantado pelos críticos estava na definição de repetência. Definiam-se como repetentes os alunos que tornavam a frequentar a mesma série por terem sido reprovados por avaliação ou frequência. Um aluno afastado por abandono ao se matricular no ano seguinte na mesma série, na mesma ou em outra escola, não era considerado repetente; aparecia como aluno novo na série. Desta forma, acabava-se superestimando a evasão e subestimando-se a repetência.

O modelo PROFLUXO, desenvolvido no Brasil por Costa-Ribeiro e Fletcher<sup>5</sup> procura corrigir essa deficiência. Trabalham com a PNAD (Pesquisa Nacional de Amostras por Domicílio), e com apenas um ano da pesquisa, conseguem identificar as taxas de promoção, repetência e evasão para todas as séries do ensino fundamental e médio. Utilizam duas perguntas do questionário da pesquisa: para quem frequenta escola, série e grau atualmente cursados; e para quem não frequenta escola, última série e grau concluídos com sucesso. O modelo identifica as transições a partir das gerações: a geração  $j-1$ , observada no ano  $t$ , representa a geração  $j$  no ano  $t+1$ . No modelo, os alunos afastados que retornam à série são contados como repetentes. Segundo os resultados obtidos pelo PROFLUXO, para a 1ª série do ensino fundamental, em 1982, as taxas de repetência e evasão eram da ordem de 52,4% e 2,3%, diferentes das taxas de 29,6% e 25,5% que eram divulgadas oficialmente<sup>6</sup>.

Klein<sup>7</sup> com os dados do Censo Escolar também propõe um modelo de fluxo escolar, buscando corrigir as distorções metodológicas a respeito da repetência. O conceito de repetência é ampliado para incluir tanto os afastados que voltam a cursar a mesma série bem como os aprovados repetentes, ou seja, alunos que embora tenham sido aprovados na série repetem a série no ano seguinte<sup>8</sup>. Atualmente, o INEP (Instituto Nacional de Estudos Pedagógicos – ligado ao Ministério da Educação e Cultura), com base no método proposto por Klein, faz as projeções de matrículas oficiais. Para o cálculo das taxas de transição, utiliza somente os dados de matrículas, aprovados e número de alunos provenientes de fora do sistema coletados nos dados dos Censos Escolares de dois anos consecutivos. Esta última informação, como será discutido mais adiante, garante que se trate adequadamente o modelo como um sistema fechado, corrigindo possíveis distorções decorrentes do retorno ao sistema regular de ensino de alunos anteriormente evadidos.

O objetivo deste trabalho é propor um modelo de fluxo escolar. A principal vantagem do modelo proposto é a de incorporar o atraso escolar nas estimativas das taxas de transição, diferentemente do modelo atual utilizado pelo INEP e também do modelo PROFLUXO, que admitem a hipótese de que as taxas de transição seguem um processo markoviano. Esta diferença

---

<sup>4</sup> Para um melhor entendimento dessa discussão, ver dentre outros, Fletcher e Ribeiro (1987), Ribeiro (1991) e Fletcher e Castro (1993).

<sup>5</sup> Para um entendimento do modelo PROFLUXO, ver Fletcher e Ribeiro (1996); para uma discussão dos resultados do modelo e comparações com outras fontes de dados ver Ribeiro (1991).

<sup>6</sup> RIBEIRO (1991), tabelas 1 e 2, pág. 9.

<sup>7</sup> Klein (1995) e Klein e Ribeiro (1991).

<sup>8</sup> Este fenômeno era mais comum para a 1ª série do EF.

apresenta-se como uma contribuição à teoria, tendo em vista a correlação positiva observada entre a repetência e a probabilidade de repetência<sup>9</sup>. Além disso, o modelo proposto utiliza a PNAD (Pesquisa Nacional de Amostras por Domicílio) do IBGE para os cálculos, o que se constitui uma vantagem frente ao modelo do INEP<sup>10</sup>, visto que a PNAD permite que se leve em conta características demográficas e sócio-econômicas nas estimativas das taxas de transição escolar. Desta forma, as taxas de transição não são constantes ao longo tempo, podendo variar em função de mudanças nas características das novas gerações. Trabalha-se com o fluxo escolar ao longo do ensino fundamental e ensino médio. A partir das taxas de transição obtidas projeta-se a PNAD com relação às características educacionais para os anos de 2004 a 2009.

O artigo está dividido em mais quatro seções além dessa introdução. A segunda seção desenvolve o modelo de fluxo proposto por essa pesquisa. A terceira seção apresenta as estimativas das taxas de transição escolar obtidas pelo modelo e a quarta seção as estimativas das matrículas. Sempre que possível são feitas comparações dos números obtidos com os divulgados pelo INEP. A quinta seção conclui.

## 2. O modelo proposto

Essa seção apresenta o modelo de fluxo de escolar proposto. A unidade de observação é a coorte escolar definida segundo os padrões oficiais<sup>11</sup>. Para identificar a transição, utiliza-se a dimensão tempo; por essa razão, trabalha-se com dois anos de PNAD: 2002 e 2003. Dessa forma, observa-se uma coorte  $j$  em  $t$  e  $t+1$ .

As hipóteses assumidas, necessárias para o cálculo das taxas de transição e projeção são:

1. As crianças que entram na escola não saem até obter, pelo menos, uma série completa.
2. A decisão de evasão é tomada no início do ano letivo.
3. Uma vez evadida, a criança nunca mais retorna à escola, ou seja, assume-se que o sistema é fechado.
4. A probabilidade de ingresso depende do atraso e das características pessoais/familiares:  $\Pr[I|A; X]$ , onde  $A$  = atraso (idade efetiva – idade ideal para estar cursando a série) e  $X$  são as características pessoais e familiares;
5. A probabilidade de evasão depende da idade, das características pessoais/familiares ( $X$ ) e do atraso escolar:  $\Pr[E|A; X; Id]$ , onde  $Id$  = idade.
6. A probabilidade de aprovação depende das características pessoais/familiares ( $X$ ), do atraso escolar e da série -  $\Pr[Ap|A; X; S]$ ,  $S$  = série escolar
7. As características ( $X$ ) são consideradas invariantes no tempo.

A primeira hipótese é necessária porque na PNAD não é possível identificar para um indivíduo que não frequenta escola atualmente, no caso de ele ter zero ano de estudo, se ele já ingressou na escola em algum momento. Portanto, indivíduos com zero ano de estudo que não frequentam escola atualmente, por hipótese, nunca frequentaram escola. A segunda hipótese será necessária para a

<sup>9</sup> Barros e Mendonça (1998) apresentam evidências de que a repetência eleva a probabilidade de repetência.

<sup>10</sup> O único corte considerado pelo INEP são os estados, isto é, as taxas de transição são estimadas separadamente para cada unidade da federação. Embora o PROFLUXO também utilize a PNAD, não se encontrou nenhum texto onde os autores tenham incorporado características sócio-demográficas nas taxas de transição estimadas.

<sup>11</sup> A idade ideal para entrar na primeira série do primeiro grau contempla todas as crianças que completam 7 anos até 31 de julho do ano  $t$ , portanto as nascidas, entre 01 de agosto do ano  $t-8$  e 31 de julho do ano  $t-7$ .

projeção. A idéia é que primeiro define-se a aprovação ou reprovação – que é conhecida apenas no final do ano letivo - e, depois disso a evasão. A hipótese três define o sistema como sendo fechado, ou seja, alunos que saíram da escola não voltam. Essa hipótese é forte (pode inclusive nos levar a superestimar a evasão escolar), principalmente para as séries mais avançadas, porém esse movimento de entrada e saída da escola não seria passível de ser identificado. Como consequência dessa hipótese, a probabilidade de ingresso só é calculada para a primeira série do primeiro grau. As taxas de transição dependem da idade ou série, ou seja, o atraso é considerado, e das características pessoais/familiares, disponíveis na PNAD – hipóteses 4, 5 e 6. Ou seja, assume-se que características das escolas ou não-observáveis não afetam as transições. Por fim, a hipótese 8 que se faz necessária na medida em que é o tempo quem identifica as taxas de transição.

Antes da apresentação da derivação das probabilidades, é importante que se faça um esclarecimento. O INEP divulga as taxas de promoção, repetência e evasão. Se o sistema de ensino é fechado, dos matriculados na série  $k$ , no ano  $t$ , parte repetirá a série em  $t+1$  (são os repetentes), parte cursará a série  $k+1$  em  $t+1$  (são os promovidos) e parte será aprovada ou reprovada e sairá da escola (são os evadidos). Logo, esses fluxos como proporção da matrícula inicial, devem somar um<sup>12</sup>. O fato é: os promovidos mais os evadidos aprovados correspondem aos aprovados; os repetentes mais os evadidos reprovados correspondem aos reprovados. No presente artigo, apresentam-se definições de aprovados (de forma que ‘1-Aprovados’ corresponde à reprovação) e evadidos e trabalha-se com esses conceitos ao longo do texto<sup>13</sup>.

### Obtendo as probabilidades

Para calcular as taxas de transição a partir da PNAD, pelo menos três procedimentos podem ser identificados<sup>14</sup>. Uma primeira idéia consiste em obter as estimativas a partir de anos sucessivos de PNAD's. Embora não se observem as mesmas crianças entre os anos, poder-se-ia acompanhar uma determinada geração com determinadas características, durante a sua trajetória escolar que permitiria gerar estimativas das probabilidades de transição. Uma outra alternativa seria a de usar apenas um ano de PNAD. Assumindo-se que o crescimento populacional esteja estável, uma coorte escolar de idade  $j$  pode ser representada um ano depois por uma coorte mais velha naquele mesmo ano. Ou seja, a identificação da transição é feita através das coortes. Este é o procedimento do PROFLUXO. E, por fim, uma terceira alternativa seria a de fixar o tempo em que uma transição ocorre - comparar apenas duas pesquisas domiciliares - e obter tais estimativas utilizando gerações sucessivas. Esta é a forma de identificação utilizada.

---

<sup>12</sup> Note que se alunos que saíram do sistema em algum momento anterior, retornam ao sistema esses fluxos serão maiores do que a matrícula da série  $k$ , no ano  $t$ . A vantagem do INEP com relação ao modelo proposto é que o INEP identifica os alunos que estão retornando ao sistema de ensino e, então, desconta esse número da matrícula inicial. Com os dados da PNAD não é possível identificar tais alunos, assim essa correção não é feita.

<sup>13</sup> Uma das vantagens do modelo proposto com relação ao profluxo está na contabilização da evasão. No profluxo não há evasão após a repetência. Para evadir, o indivíduo tem que ser aprovado. No modelo proposto esta restrição é imposta apenas para a 1ª série do ensino fundamental, como já mencionado anteriormente.

<sup>14</sup> Leon e Menezes-Filho (2002) calculam as taxas de transição escolares a partir dos dados da Pesquisa Mensal de Emprego – PME–IBGE - para as séries diplomas. Como a PME acompanha os mesmos indivíduos por um determinado tempo, a transição é identificada a partir dos indivíduos.

## - As probabilidades de ingresso

Seja  $F_{0j}$  o número de crianças fora da escola e sem nenhuma escolaridade no ano com idade ideal para cursar a série  $j$ , então a probabilidade de ingresso condicional ao atraso pode ser definida como

$$\Pr[I|A=j] = 1 - \frac{F_{0j+1}}{F_{0j}}.$$

Pense no caso da geração que tem idade ideal para ingressar na 1ª série do ensino fundamental (EF) em 2003 – geração das crianças que completam 7 anos até 31/julho/2003. Para essa geração, exclusivamente,  $F_{0j}$  é o número de crianças da geração - visto que por definição todas estão fora da escola em 2002 com zero ano de estudo; e  $F_{0j+1}$  é o número de crianças dessa geração que permaneceu fora da escola em 2003. Ou seja, a razão entre  $F_{0j+1}$  e  $F_{0j}$  é exatamente a proporção da geração que não ingressou; logo ‘1- razão’ é a proporção da geração que ingressou.

Como as características sócio-demográficas das crianças são consideradas na estimação das probabilidades, não se tornaram viáveis – em razão do número de observações – procedimentos não-paramétricos para os cálculos, levando ao uso de técnicas paramétricas. Definindo  $N$  como o tamanho da geração, chega-se que

$$\frac{F_{0j+1}}{F_{0j}} \frac{N}{N} = \frac{\Pr[F=0|Id=j+1]}{\Pr[F=0|Id=j]} \quad \text{onde, } \Pr[F=i|Id=j] \text{ é a}$$

probabilidade do indivíduo estar fora da escola com  $i$  anos completos de estudo condicional a ter idade ideal para cursar a  $j$ -ésima série. Assim,

$$\Pr[I|A=j] = 1 - \frac{\Pr[F=0|Id=j+1]}{\Pr[F=0|Id=j]} \quad (1)$$

Note que como apenas duas PNAD's serão utilizadas, cada probabilidade de ingresso condicional ao atraso será obtida por uma geração diferente. O procedimento a ser adotado é o que segue.

- 1) Seleciona-se todas gerações com idade para estar na escola no período 1 e possuem crianças fora da escola e sem nenhuma escolaridade no período 0.
- 2) Excluem-se as gerações cuja probabilidade de ingresso pode ser considerada igual a 0.
- 3) Define-se  $\Pr[F=0] = f(Z)$ . Assume-se que  $f(z)$  comporta-se como uma logística e que  $Z = \alpha + X\beta + \gamma T + G\eta + T.G\lambda$ , onde:

- $X$  = vetor de características pessoais/familiares que incluem uma *dummy* para gênero, uma *dummy* para o fato de a criança residir em área urbana e uma para o fato de residir em área metropolitana, quatro *dummies* para captar a escolaridade dos pais – que foi considerada a maior entre o pai e a mãe (possibilidades: 0 ano de estudo; de 1 a 4 anos de estudo; de 5 a 8; de 9 a 11; e 12 ou mais); e um conjunto de *dummies* para as uf's do país.
- $T$  = variável *dummy* que é igual a 1 quando a PNAD considerada refere-se ao ano de 2003; alternativamente, utiliza-se  $T=1$  como referência ao período 1.
- $G$  = conjunto de variáveis *dummies* para  $n-1$  gerações consideradas.
- $T.G$  = interação entre a *dummy* de tempo e as *dummies* de geração; esta interação mostra que o efeito tempo age de forma diferente para cada geração.

Então, considerando  $G = i$  a geração cujas crianças que estão ingressando na escola no período 1 possuem  $i$  anos de atraso, tem-se:

$$\Pr[I|A=i; X] = 1 - \frac{\Pr[F=0|T=1; G=i; X]}{\Pr[F=0|T=0; G=i; X]} \quad (2)$$

Em suma, o procedimento consiste em juntar todas as gerações e as duas PNAD's e estimar uma única regressão para  $\Pr[F = 0] = f(Z)$ . É o *pooling* de gerações nas estimativas que permite que se obtenha probabilidades condicionais ao atraso. Para as estimativas foram selecionadas 6 gerações, de forma que se estima a probabilidade de ingresso condicional a ter até 5 anos de atraso – aproximadamente, estar ingressando na 1ª série do EF com 12 anos de idade. Essa última probabilidade foi imputada para os adolescentes com idade entre 13 e 16 anos e para o restante dos casos, imputou-se uma probabilidade de ingresso igual a 0.

Duas observações devem ser feitas: i) muitas das crianças ingressam antes de atingir a ‘idade ideal’ de ingresso da forma como definida. A fim de neutralizar esse impacto nas estimativas, considerou-se que essas crianças ainda estavam fora da escola; ii) como, por definição, todas as crianças com idade ideal para estar na 1ª série do EF estão fora da escola em 2002, estimou-se uma equação à parte para a probabilidade de estar fora da escola para esta geração, considerando apenas o ano de 2003. Note também que para esta geração  $\Pr[F = 0 | T = 0; G = i; X] = 1$ .

### As probabilidades de evasão

Seja  $Y_{ij}$  o número de crianças que possui idade ideal para cursar série  $j$ , mas está cursando a série  $i$  e  $F_{ij}$  a proporção de crianças com  $i > 0$  anos de estudo e fora da escola no ano  $j$ , o ano cuja idade é a

$$\text{ideal para se cursar a } j\text{-ésima série. Então, } \Pr[E | Id = j; Ed = i] = \frac{F_{ij} - F_{ij-1}}{Y_{i+1j} + F_{ij} - F_{ij-1}} \text{ ou}$$

$$\Pr[E | Id = j; Ed = i] = \frac{\Pr[F = i | Id = j] - \Pr[F = i | Id = j - 1]}{\Pr[S = i + 1 | Id = j] + \Pr[F = i | Id = j] - \Pr[F = i | Id = j - 1]} \quad (3)$$

Nas equações acima,  $Ed = i$  significa que o indivíduo possui  $i$  anos completos de estudo, enquanto  $S = i$  significa que ele está cursando a série  $i$ . Ou seja, a probabilidade da criança evadir condicional a ter  $i$  anos de estudo completos de estudo e possuir idade ideal para estar cursando a série  $j$  no ano  $j$  é dada pela razão entre o número de crianças com essas características que está fora da escola no ano  $j$  (ou seja, que evadiu no ano  $j$ ) e o número de crianças com tais características que poderia evadir naquele ano. Para que se conte apenas as crianças com essas características que evadiram ou que poderiam evadir no ano  $j$  é preciso desconsiderar aquelas crianças fora da escola hoje mas que já tinham evadido antes, isto é,  $F_{ij-1}$  (número de crianças que já estava fora da escola no momento  $j-1$ , com  $i$  anos de estudo, e que tinha idade ideal para estar cursando a série  $j-1$  no período anterior). O termo  $\Pr[S = i + 1 | Id = j] + \Pr[F = i | Id = j]$  por outro lado corresponde a  $\Pr[Ed = i | Id = j]$ , que é simplesmente a probabilidade de ter  $i$  anos completos de estudo (podem ser alunos matriculados na série ou fora da escola), condicional a ter idade ideal para cursar a série  $j$  no ano  $j$ . Então,

$$\Pr[E | Id = j; A = j - i - 1] = \frac{\Pr[F = i | Id = j] - \Pr[F = i | Id = j - 1]}{\Pr[Ed = i | Id = j] - \Pr[F = i | Id = j - 1]} \quad (4)$$

Diferentemente do caso anterior, nesse caso será necessário estimar diversas regressões: duas para cada  $Ed = i$ . Considere o caso onde  $i = 1$ , ou seja, o caso de evasão com um ano completo de estudo.

- 1) Seleciona-se todas as gerações que, no período 1, possuam pessoas fora da escola com 1 ano completo de estudo e que possuam crianças na 2ª série.
- 2) Definem-se  $\Pr[F = 1] = f(Z_1)$  e  $\Pr[Ed = 1] = f(Z_2)$ , onde

$$Z_i = \alpha_i + X\beta_i + \gamma_i T + G\eta_i + T.G\lambda_i \quad (i = 1, 2).$$

Assume-se que estas probabilidades se comportam como uma logística e estimam-se as probabilidades de estar fora da escola com 1 ano completo de estudo e de ter um ano completo de estudo. O vetor de características  $X$  é o mesmo da equação de ingresso. Novamente, é o *pooling* de gerações nas estimativas que permite que se obtenha as probabilidades condicionais ao atraso.

Como no caso anterior, juntam-se todas as gerações e as duas PNAD's e estimam-se as duas regressões acima, onde  $G = i$  é a geração cujas crianças com  $Ed = 1$  no período 1 possuem 0 ano de atraso. Isso nos fornece as seguintes probabilidades condicionais.

$$\Pr[E|Id=2; A=0; X] = \frac{\Pr[F=1|T=1; G=0; X]}{\Pr[Ed=1|T=1, G=0; X]} \quad (5.a)$$

$$\Pr[E|Id=3; A=1; X] = \frac{\Pr[F=1|T=1; G=1; X] - \Pr[F=1|T=0; G=1; X]}{\Pr[Ed=1|T=1, G=1; X] - \Pr[F=1|T=0; G=1; X]} \quad (5.b)$$

$$\Pr[E|Id=4; A=2; X] = \frac{\Pr[F=1|T=1; G=2; X] - \Pr[F=1|T=0; G=2; X]}{\Pr[Ed=1|T=1, G=2; X] - \Pr[F=1|T=0; G=2; X]} \quad (5.c)$$

etc.

O mesmo procedimento pode ser aplicado para  $Ed > 1$  e, assim, obter as demais probabilidades. No caso de  $Ed = 2$ , obtém-se  $\Pr[E|Id=3; A=0; X]$ ,  $\Pr[E|Id=4; A=1; X]$ ,  $\Pr[E|Id=5; A=2; X]$  etc. Trabalha-se com 6 gerações, de forma que teremos a probabilidade de evasão para cada  $i$  com até 6 anos de atraso. Para as demais gerações imputou-se a probabilidade de evasão da última geração.

Algumas observações devem ser feitas. A primeira delas é, novamente, com relação ao tratamento dado aos adiantados. Neste caso, a opção foi excluí-los das estimativas. Ao se corrigir os 'adiantados' acabou-se superestimando a evasão para algumas séries - por exemplo, para a sétima série do EF. Um segundo ponto é que as estimativas para a probabilidade de evasão condicional a ter atraso igual a zero foram feitas separadamente. Para este caso, em particular,  $\Pr[F=i|Id=j-1]=0$ , por definição, por que quem no ano  $j$  está na série correta, independente de vir a evadir ou não no ano  $j$ , não poderia ter evadido no ano anterior com  $i$  anos de estudo visto que certamente tinha  $i-1$  anos de estudo. Além disso, as probabilidades de estar fora da escola condicional a ter zero ano de atraso, para todos os anos de estudo, foram obtidas a partir de uma única equação. Isto porque o número de crianças que está fora da escola com potencial para estar na série correta é bastante pequeno, inviabilizando estimativas individuais para cada  $i$  ano de estudo.

### As probabilidades de aprovação

Por fim, para se obter as probabilidades de aprovação, seria conveniente definir a probabilidade de avanço como  $\Pr[Av|Ed=i]$ , que é a probabilidade de se ter  $i+1$  anos completos de estudo no ano 1, condicional ter  $i$  anos de estudo no ano 0. Defina também  $\Pr[S=i|T=0,1]$  como a probabilidade de se estar cursando a série  $i$  no tempo  $T = t$ . De posse dessas probabilidades para as diversas gerações, pode-se obter todas as probabilidades de aprovação desejadas. A probabilidade de aprovação da série  $j$  para a série  $j+1$  é razão entre os que ganharam um ano a mais de estudo (produto entre as probabilidades de ter  $i$  anos de estudo em  $t=0$  e avançar) e os que estavam freqüentando a série  $j$  no ano 0.

Veja o caso da geração que no período  $t=1$  possui idade ideal para cursar a 5ª série, a qual denomina-se de geração  $G = 5$ . Para essa geração pode-se obter as seguintes probabilidades de aprovação:

$$\Pr[Ap | A=0; S=4; X] = \frac{\Pr[Ed=3 | T=0; G=5; X] \Pr[Av | Ed=3; G=5; X]}{\Pr[S=4 | T=0; G=5; X]} \quad (6.a)$$

$$\Pr[Ap | A=1; S=3; X] = \frac{\Pr[Ed=2 | T=0; G=5; X] \Pr[Av | Ed=2; G=5; X]}{\Pr[S=3 | T=0; G=5; X]} \quad (6.b)$$

$$\Pr[Ap | A=2; S=2; X] = \frac{\Pr[Ed=1 | T=0; G=5; X] \Pr[Av | Ed=1; G=5; X]}{\Pr[S=2 | T=0; G=5; X]} \quad (6.c)$$

$$\Pr[Ap | A=3; S=1; X] = \frac{\Pr[Ed=0 | T=0; G=5; X] \Pr[Av | Ed=0; G=5; X]}{\Pr[S=1 | T=0; G=5; X]} \quad (6.d)$$

Por sua vez, as probabilidades de avanço são:

$$\Pr[Av | Ed=3; G=5; X] = \frac{\Pr[Ed=4 | T=1; G=5; X]}{\Pr[Ed=3 | T=0; G=5; X]} \quad (7.a)$$

$$\Pr[Av | Ed=2; G=5; X] = \frac{\Pr[Ed=3 | T=1; \dots] - \Pr[Ed=3 | T=0; \dots] \{1 - \Pr[Av | Ed=3; \dots]\}}{\Pr[Ed=2 | T=0; \dots]} \quad (7.b)$$

$$\Pr[Av | Ed=1; G=5; X] = \frac{\Pr[Ed=2 | T=1; \dots] - \Pr[Ed=2 | T=0; \dots] \{1 - \Pr[Av | Ed=2; \dots]\}}{\Pr[Ed=1 | T=0; \dots]} \quad (7.c)$$

$$\Pr[Av | Ed=0; G=5; X] = \frac{\Pr[Ed=1 | T=1; \dots] - \Pr[Ed=1 | T=0; \dots] \{1 - \Pr[Av | Ed=1; \dots]\}}{\Pr[Ed=0 | T=0; \dots]} \quad (7.d)$$

Tome como exemplo o segundo caso, o caso em que a criança tem um ano de atraso, ou seja, ela está na terceira série - mas tem idade ideal para estar na 4ª série - com 2 anos de estudo. O objetivo é calcular para essa criança a probabilidade dela ganhar um ano a mais de estudo, isto é, a probabilidade de ela ir para a 4ª série e então passar a ter 3 anos de estudo. A equação de avanço é a razão entre a probabilidade de a criança ter 3 anos de estudo no ano  $t=1$  e a probabilidade de ela ter 2 anos de estudo no ano  $t=0$ . Note entretanto, que existem crianças desta geração que já tem 3 anos de estudo em  $t=0$ , sendo que parte dessas irá ser aprovada e ganhar um ano a mais de estudo e parte será reprovada e permanecerá com 3 anos de estudo. Logo para calcular o avanço de 2 para 3 anos de estudo é preciso descontar essa parcela de crianças que já tem 3 anos e que vai ser reprovada entre  $t=0$  e  $t=1$ . Este desconto pode ser calculado como sendo a probabilidade de ter 3 anos de estudo em  $t=0$  multiplicada pela probabilidade de não avançar para 4 anos de estudo (termo entre chaves). Note o uso recursivo das equações: para calcular a probabilidade de avanço para 3 anos de estudo, utiliza-se a probabilidade de avanço para 4 anos de estudo. O procedimento realizado para geração  $G = 5$  é repetido para as demais gerações relevantes e, assim, obtêm-se as demais probabilidades de aprovação requeridas.

Note que para o 3º colegial, a probabilidade de aprovação corresponde à saída do sistema. Note ainda que o cálculo dessa probabilidade de aprovação condicional a ter atraso positivo envolve termos como, por exemplo,  $\Pr[Av | Ed=11; G=Y; X]$ , ou seja, que seria o equivalente à probabilidade avançar de 11 anos de estudo para 12 anos de estudo, escolaridades relativas ao ensino superior.



Como não se modela o ensino superior, considerou-se essa probabilidade, bem como as superiores a essa, como sendo iguais a zero.

Para a obtenção das probabilidades de avanço, nenhuma nova estimação é necessária, uma vez que as probabilidades de se possuir  $i$  anos completos de estudo já foram obtidas no procedimento anterior (probabilidades de evasão)<sup>15</sup>. Para obtenção das probabilidades de aprovação precisa-se apenas das probabilidades de se cursar a série  $i$ . Para se obter  $\Pr[S=i]$  o procedimento é similar ao anterior: estima-se uma regressão para cada  $i$ . Neste caso, no entanto, a estimação pode ser feita apenas com a PNAD do ano 0. Então, estima-se  $\Pr[S=i] = f(Z)$ , onde  $Z = \alpha + X\beta + G\eta$ . Também para essas probabilidades, pelos motivos apresentados anteriormente, os adiantados foram excluídos da estimativa.

### A expansão

Para se projetar a PNAD para o futuro, no que se refere às características educacionais, necessita-se saber, para cada pessoa amostrada, suas atuais situações educacionais e suas probabilidades de transição. Isto é, para:

- crianças fora da escola com zero anos de estudo no ano  $t-1$  estima-se a probabilidade de ingresso: essas crianças podem ingressar na escola ou não no período  $t$ ;
- crianças que freqüentam escola e estão na série  $k$  no ano  $t-1$  estima-se primeiramente uma probabilidade de aprovação: elas podem ser aprovadas e freqüentar a série  $k+1$  no período  $t$  ou podem ser reprovadas e voltar a freqüentar a série  $k$  no período  $t$ . Decidido a aprovação, imputa-se a probabilidade de evasão. Assim, no início do período  $t$ , essas crianças decidirão se freqüentarão a série mesmo ou se irão evadir. Assim, parte das crianças aprovadas e parte das reprovadas evadem. Exceção para as crianças que freqüentam a 1ª série do EF, as quais só sairão da escola se forem aprovadas.

Note, entretanto, que as probabilidades são números definidos entre  $(0,1)$ . Como decidir se a criança ingressa ou não quando sua probabilidade de ingresso é 0,65? Para traçar o futuro das crianças, com base nas probabilidades estimadas, recorreu-se a um sorteio. Por exemplo, cada criança tem sua probabilidade de ingressar. Cria-se também para cada criança, fora da escola com probabilidade de ingressar, um número aleatório, a partir de uma distribuição uniforme no intervalo  $[0,1]$ . Se a probabilidade de ingresso da criança superar ou empatar com o número aleatório, dizemos que essa criança ingressou; caso contrário, dizemos que permaneceu fora da escola. O mesmo vale para as demais probabilidades.

## 3. Probabilidades de Transição

Essa seção apresenta comparações de alguns indicadores escolares obtidos a partir do modelo proposto com os números do INEP. A tabela 1 abaixo apresenta os indicadores de atendimento escolar<sup>16</sup> e escolarização líquida<sup>17</sup>.

<sup>15</sup> Na realidade necessitamos acrescentar às estimativas da seção anterior o caso onde  $Ed = 0$ .

<sup>16</sup> A taxa de 'atendimento escolar' é a razão entre o número de indivíduos que freqüentam escola de determinada faixa etária e o total da população dessa determinada faixa etária.

Tabela 1: Atendimento Escolar e Escolarização Líquida

ANO	Atendimento Escolar		Escolarização Líquida	
	EF	EM	EF	EM
2000 (censo escolar)	96,40	83,00	94,30	33,30
2003 (modelo proposto)	95,75	82,83	95,72	45,54
Projeções				
2004	95,94	83,34	95,90	48,69
2005	95,76	84,44	95,75	49,85
2006	95,60	84,48	95,59	49,25
2007	95,37	85,12	95,37	48,17
2008	95,35	85,15	95,35	48,30
2009	95,37	85,70	95,36	47,87

Embora os anos sejam diferentes, os números obtidos pelo modelo proposto são bastante próximos aos do INEP. A maior diferença é com relação à escolarização líquida para o ensino médio, uma diferença de 12 p.p. Provavelmente, os números obtidos devem estar incorporando o avanço no ensino médio que vem ocorrendo nos últimos anos, o que é confirmado em recente matéria publicada no ‘Informativo do INEP’. Segundo o informativo, a escolarização líquida para o ensino médio em 2003 era dez pontos maior do que a registrada no Brasil em 2000<sup>17</sup>. É interessante notar que de acordo com as projeções realizadas, esse número deve crescer até 2006 alcançando quase 50%, estabilizando-se em seguida em torno dos 48%.

A tabela 2 a seguir apresenta as taxas de transição: aprovação e evasão. Primeiramente, é importante enfatizar que os períodos de análise são distintos e assim os números não são diretamente comparáveis, principalmente se políticas educacionais que interferem nas transições – como, por exemplo, a ‘progressão automática’ ou ‘progressão continuada’ – foram adotadas no período. Em termos da aprovação, em geral, os números obtidos são inferiores aos do INEP. As exceções ficam para a 1ª e 7ª séries do ensino fundamental e 1º e 2º colegial, quando os números obtidos são maiores e para 5ª e 6ª séries do ensino fundamental quando há igualdade entre as taxas. As maiores diferenças ficaram para o 1º e 3º colegial: no 1º colegial a taxa obtida é 11 p.p. maior e no 3º 14 p.p. menor. Uma diferença razoável também aparece na 4ª série do ensino fundamental (9 p.p. a favor das taxas via censo escolar). Com relação à projeção, de forma geral, para o ensino fundamental, há dois movimentos: primeiramente um aumento das probabilidades entre 2003 e 2004 e, então, observa-se uma diminuição das taxas de aprovação ao longo do período. Para as séries do ensino médio, ao contrário, há uma tendência de redução da taxa de aprovação ao longo do tempo. A redução é de 5 p.p., 11 p.p. e 12 p.p., respectivamente, para o 1º, 2º e 3º colegial, entre 2004 e 2009.

<sup>17</sup> A taxa de ‘escolarização líquida’ é a razão entre a matrícula no nível de ensino k de indivíduos pertencentes à faixa etária i teoricamente adequada ao nível de ensino k e a população da faixa etária i teoricamente adequada ao nível de ensino k.

<sup>18</sup> Os dados são da própria PNAD de 2003. Para maiores detalhes, consultar: <http://www.inep.gov.br/informativo/informativo94.htm>

Tabela 2: Taxas de aprovação e evasão

ANO/Série	Ensino Fundamental								Ensino Médio		
	1 <sup>a</sup>	2 <sup>a</sup>	3 <sup>a</sup>	4 <sup>a</sup>	5 <sup>a</sup>	6 <sup>a</sup>	7 <sup>a</sup>	8 <sup>a</sup>	1 <sup>a</sup>	2 <sup>a</sup>	3 <sup>a</sup>
<b>Aprovação (%)</b>											
2000-2001 (censo escolar)	67,9	78,8	82,0	84,1	73,8	79,7	80,0	82,6	68,0	76,4	84,8
2003 (nosso)	72,88	76,62	76,19	74,94	73,86	79,05	86,56	80,26	79,12	77,99	70,79
<b>Projeção</b>											
2004	74,64	77,91	78,22	77,31	75,02	80,15	85,59	81,76	79,65	77,95	73,90
2005	74,51	77,78	75,83	75,76	73,41	79,27	84,51	80,93	78,90	76,27	72,13
2006	74,20	78,06	74,65	77,53	74,39	79,38	84,67	81,16	77,13	73,46	70,00
2007	72,60	77,14	75,28	77,06	73,45	78,84	84,51	80,23	76,48	70,96	66,07
2008	72,10	77,72	75,63	77,47	73,51	79,56	84,38	80,67	75,76	68,24	63,55
2009	71,66	78,00	75,34	77,41	73,50	79,68	84,09	80,18	74,87	66,52	62,09
<b>Evasão (%)</b>											
2000 (censo escolar)	1,0	3,1	4,0	5,2	6,9	8,5	6,5	11,4	11,4	7,0	2,6
2003 (nosso)	0,41	0,67	1,00	2,41	4,60	4,60	4,80	7,50	6,43	8,48	3,53
<b>Projeção</b>											
2004	0,58	0,62	0,78	2,41	4,15	4,31	4,58	7,33	6,75	7,76	2,56
2005	0,37	0,37	0,80	2,10	3,85	3,81	5,62	6,56	5,80	6,53	2,12
2006	0,50	0,58	0,93	2,57	3,64	3,80	5,18	6,56	5,80	6,53	2,12
2007	0,60	0,64	1,04	2,25	3,66	3,88	5,10	6,55	6,40	5,93	2,04
2008	0,54	0,66	0,97	1,77	3,70	3,93	5,40	6,45	6,15	7,38	3,14
2009	0,55	0,43	1,03	2,21	3,80	3,63	4,11	6,29	6,09	7,09	3,58

As taxas de evasão obtidas são sistematicamente inferiores às do Censo Escolar, com exceção das estimadas para as duas últimas séries do ensino médio, quando os números obtidos apresentaram-se maiores (em 1,5 p.p. e 1 p.p., respectivamente, para o 2º e 3º colegial). Novamente, quanto à projeção, observa-se um aumento das taxas de evasão entre 2003 e 2004, seguido de uma tendência de redução das taxas, na maior parte dos casos (em 8 dos 11 possíveis).

Em resumo, os fluxos apontam uma menor evasão acompanhada de uma menor aprovação, algo razoável, quando se acredita na hipótese de que os indivíduos que evadem são aqueles menos preparados e, portanto, os que possuem maior probabilidade de reprovação. Como as taxas de transição são condicionais às características sócio-demográficas dos alunos, uma ‘piora’ desse perfil será captado pelas taxas de transição.

Uma das grandes vantagens do modelo proposto está no fato das taxas de transição estarem condicionadas ao atraso e às características individuais dos alunos. A seguir apresentam-se os índices calculados por atraso e por escolaridade dos pais.

Primeiramente a probabilidade de ingresso na 1ª série do ensino fundamental. Como não há grande variabilidade ao longo dos anos nesse índice - o desvio-padrão das estimativas ao longo dos cinco anos de análise para cada nível de atraso não atinge 1 ponto percentual – optou-se por apresentar apenas os valores iniciais e finais. Observe, como esperado, que o ingresso diminui monotonicamente com o atraso e, mais interessante, que o ingresso no ensino fundamental está praticamente universalizado visto que, de acordo com as taxas obtidas, até 3 anos após a idade ideal de ingresso, todas as crianças já estarão na escola.

Tabela 3  
Probabilidade de Ingresso na 1ª Série do ensino fundamental segundo o atraso escolar (%)

Probabilidade de Ingresso (%)	Atraso (medido em anos)					
	0	1	2	3	4	5
2004	86,95	70,15	54,89	35,23	35,33	35,11
2009	86,58	70,13	54,34	35,59	35,69	34,98
Desvio-Padrão (2004-2009)	0,46	0,75	0,27	0,59	0,49	0,58

Na tabela 4, estão as probabilidades de aprovação e de evasão por níveis de atraso. Os números apresentados são médias simples das probabilidades obtidas para cada uma das séries do ensino fundamental e ensino médio. Como na maior parte dos casos, a variabilidade das taxas de transição ao longo dos anos para cada nível de atraso não ultrapassou 2 p.p., optou-se novamente por apresentar apenas os valores iniciais e os finais.

A aprovação diminui monotonicamente com o atraso. Em 2004, para os alunos sem atraso escolar a probabilidade de ser aprovado era de 90% aproximadamente; para os que tinham três anos de atraso escolar, tal probabilidade caía em 30%; e para os que tinham seis anos de atraso, caía em 60%. Esses números reforçam a importância de se considerar o atraso nas estimativas das taxas de transição. Note que existe uma tendência de queda expressiva das probabilidades de aprovação, para os mais atrasados (cinco e seis anos de atraso). Uma possível explicação está justamente nas probabilidades de evasão apresentadas em seguida.

Tabela 4  
Probabilidades de Aprovação e Evasão segundo o atraso escolar (%) – Média do EF e EM

Taxas de transição	Atraso (medido em anos)						
	0	1	2	3	4	5	6
<b>Aprovação (%)</b>							
2004	89,84	79,74	69,56	59,87	51,60	38,98	32,96
2009	91,41	77,05	68,49	59,25	49,05	28,15	23,59
Desvio-Padrão (2004-2009)	0,56	0,91	0,34	1,35	1,98	4,15	3,80
<b>Evasão (%)</b>							
2004	1,19	2,58	5,04	7,19	9,77	13,08	18,6
2009	1,24	2	4,05	7,16	7,08	9,47	12,44
Desvio-Padrão (2004-2009)	0,05	0,24	0,40	0,08	0,88	1,47	2,54

A taxa de evasão cresce com o atraso escolar, de 1% para os alunos sem atraso para 19%, em 2004, para os alunos com seis anos de atraso. Entretanto, enquanto que as taxas de evasão até três anos de atraso ficam, segundo as estimativas obtidas, praticamente constantes no período, para níveis de atraso superiores a quatro anos, principalmente superiores a cinco anos de atraso, essas apresentam uma tendência forte de redução em 2009 relativamente a 2004. Novamente, se for razoável a hipótese de que os alunos que evadem são em geral aqueles menos preparados, essa seria uma possível explicação para a tendência de queda da probabilidade de aprovação para 2009, apresentada anteriormente.

Uma outra análise interessante é confrontar os indicadores estimados segundo o nível de escolaridade dos pais. Se o background familiar é um fator importante para explicar a escolaridade dos filhos<sup>19</sup>, as taxas de transição escolar devem refletir tais diferenças.

Tabela 5  
Probabilidade de ingresso na idade correta (sem atraso) segundo a escolaridade dos pais (%)  
Média do EF e EM

<b>Ingresso sem atraso (%)</b>	Escolaridade dos pais				
	Sem instrução	1 a 4	5 a 8	9 a 11	12 ou mais
2004	69,69	82,96	87,04	91,01	93,53
2009	67,65	79,76	86,87	90,05	92,65
Desvio-Padrão (2004-2009)	0,91	1,37	0,45	0,60	0,83

As estimativas são para o ingresso na idade correta, ou seja, condicional a não ter atraso escolar. Como esperado, quanto maior a escolaridade dos pais, maiores são as taxas de ingresso: para filhos de pais sem instrução a probabilidade média de ingresso é de 70%; já para filhos de pais com 2º grau completo ou mais, a taxa é de 93,5%. Novamente, há pouca variação das taxas de ingresso ao longo do tempo, por isso a opção de apresentar apenas os valores inicial e final. Entre 2004 e 2009, as taxas declinam levemente, com uma queda máxima de 3 p.p. para filhos de pais que possuem entre 1 e 4 anos de estudo.

Tabela 6  
Probabilidades de Aprovação e de Evasão segundo a escolaridade dos pais (%)  
Média do EF e EM

Taxas / Escolaridade dos pais	Sem instrução	1 a 4	5 a 8	9 a 11	12 ou mais
<b>Aprovação (%)</b>					
2004	69,93	75,05	81,08	80,62	76,36
2005	69,98	74,21	80,51	78,72	73,71
2006	70,72	74,42	80,57	78,37	71,07
2007	68,85	74,00	79,89	77,16	68,86
2008	69,72	74,24	79,78	77,08	66,47
2009	69,88	74,87	79,68	76,73	63,63
<b>Evasão (%)</b>					
2004	7,12	5,81	4,11	2,41	1,34
2005	5,92	5,1	3,75	2,28	1,55
2006	6,10	4,9	3,82	2,34	1,45
2007	5,55	4,93	3,76	2,43	1,62
2008	5,24	4,87	3,76	2,74	1,94
2009	4,27	4,81	3,73	2,63	1,97

<sup>19</sup> Evidências sugerem que crianças com background familiar mais pobre tendem a abandonar o sistema escolar mais cedo [Filmer, D. and Pritchett, L. (1998)]. Além disso, o relatório Coleman, como ficou conhecido, e estudos de autores

A probabilidade de aprovação como função da escolaridade dos pais comporta-se como uma função quadrática, ou seja, inicialmente cresce (até atingir a escolaridade de 5 a 8 anos de estudo) e depois decresce suavemente. Uma possível explicação pode estar na qualidade da escola, algo não captado nas estimativas: se filhos de pais mais educados estudam em escolas de melhor qualidade e mais rígidas quanto à aprovação, a probabilidade de aprovação dessas crianças poderia ser menor. Quanto à projeção, chama atenção a queda das taxas de aprovação das crianças que são filhas de pais mais educados (com 12 anos ou mais de escolaridade): de 76% em 2004 para 64% em 2009. A explicação neste caso não está nas taxas de evasão, as quais inclusive crescem levemente no período. Para a evasão, o resultado é o esperado: quanto maior a escolaridade dos pais, menor é a probabilidade de evasão das crianças. Quanto às projeções, destaque deve ser dado para a queda da probabilidade de evasão das crianças cujos pais possuem um nível menor de escolaridade.

#### 4. Matrículas

Uma primeira análise é comparar os dados obtidos para as matrículas em 2003 e 2004 com os divulgados pelo INEP com base nos dados do Censo Escolar.

Tabela 7  
Matrículas no Ensino Fundamental e Ensino Médio

	Ensino Fundamental*		Ensino Médio*	
	Censo Escolar**	Estimativas Obtidas***	Censo Escolar**	Estimativas Obtidas***
2003	37.754.636	31.999.870	10.160.491	8.911.425
2004	38.021.442	31.274.722	10.878.410	9.208.253
Crescimento entre 2003 e 2004 (%)	0,707	-2,266	7,066	3,331

\* Inclui Ensino Regular e Ensino de Supletivo.

\*\*Os números são os divulgados pelo INEP com base no Censo Escolar do respectivo ano.

\*\*\*Para 2003, os valores são os observados na PNAD de 2003. Para 2004, são os números obtidos pelo modelo de fluxo escolar proposto.

A tabela 7 apresenta os dados das matrículas de ensino fundamental e médio – considerando os alunos matriculados no ensino regular e no supletivo – com base nas informações do Censo Escolar e da PNAD, sendo que no caso da PNAD, os dados de 2003 são os observados e os de 2004 são os obtidos a partir do modelo fluxo escolar proposto. Como é possível observar, existe uma diferença entre os números da PNAD e os do Censo Escolar: as matrículas a partir do Censo Escolar são 18% e 14% maiores do que as da PNAD, em 2003, no ensino fundamental e médio, respectivamente. Parte da diferença deve-se ao fato da PNAD não incluir em sua base amostral a região rural dos estados do norte do país. Como, segundo os dados do Censo Demográfico 2000, nessa área residem apenas 2% do total da população brasileira, deve haver outra razão para essa subestimação. Uma outra explicação pode estar nos próprios dados do Censo Escolar. As informações do Censo são preenchidas pelas escolas e após a implementação do FUNDEF surgiram denúncias de superestimação das matrículas. Isto porque as transferências do FUNDEF são

---

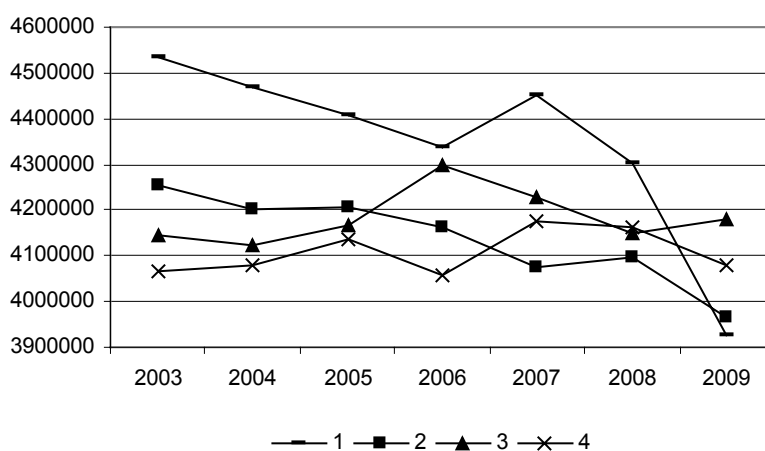
como Hanusheck (1986, 1989) apontam que o principal determinante do desempenho escolar dos alunos são as características de background familiar.

proporcionais às matrículas<sup>20</sup>. Assim, parte da diferença pode estar na superestimação dos dados do Censo Escolar.

Com relação à tendência entre os anos, tem-se que: em se tratando do ensino fundamental, as matrículas ficam praticamente constantes entre 2003 e 2004, de acordo com os dados do INEP e apresentam uma redução de 2,3% de acordo com as estimativas obtidas. Em se tratando do ensino médio, as matrículas crescem, entre 2003 e 2004, segundo as duas fontes, mas em maior magnitude segundo as estimativas do INEP (7,1% contra 3,3%).

Os gráficos a seguir apresentam as matrículas escolares por séries para cada um dos anos de expansão. Primeiramente com relação ao gráfico 1. As estimativas apontam para uma queda significativa da matrícula para a primeira série do ensino fundamental: de mais de 4,5 milhões para menos de 4 milhões. Provavelmente, deve estar havendo um ajuste das matrículas nas séries iniciais em decorrência do processo de universalização do acesso a escola que ocorreu na década passada: as matrículas nas primeiras séries estão se normalizando, aproximando-se das medidas de tamanho populacionais das coortes. Para a segunda série o comportamento é análogo embora mais suave. As matrículas das terceira e quarta séries inicialmente caminham juntas, crescendo suavemente, de acordo com o processo de ajuste indicado anteriormente. Entre 2005 e 2006, há uma quebra: enquanto que a matrícula da terceira série aumenta, a da quarta série diminui. O movimento oposto aparece entre 2006 e 2007. Entre 2007 e 2008, as matrículas nas duas séries ficam bastante próximas, separando-se novamente entre 2008 e 2009.

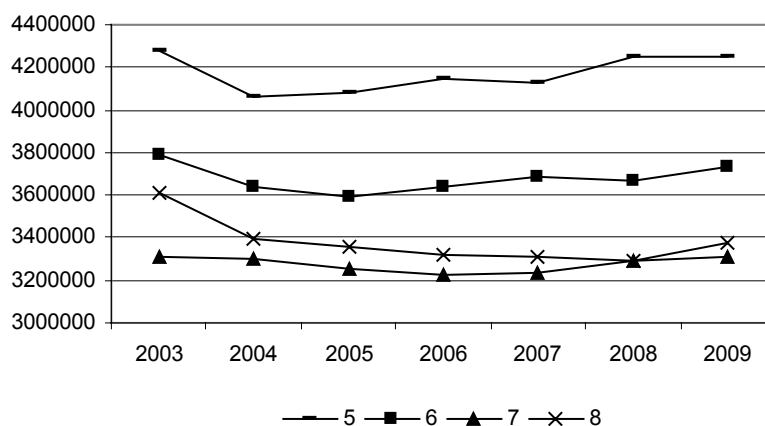
Gráfico 1: Projeção de matrículas para a 1 a 4ª série do Ensino Fundamental



As matrículas de 5 a 8ª série apresentam trajetórias bem mais suaves. Com exceção da 7ª série, para a qual a matrícula estimada em 2009 é inferior a de 2004 (uma diferença de 200.000 alunos aproximadamente), para as demais séries, há uma leve queda do número de matriculados entre os anos, voltando a crescer a partir de 2008, de forma que os valores inicial e final são bem parecidos.

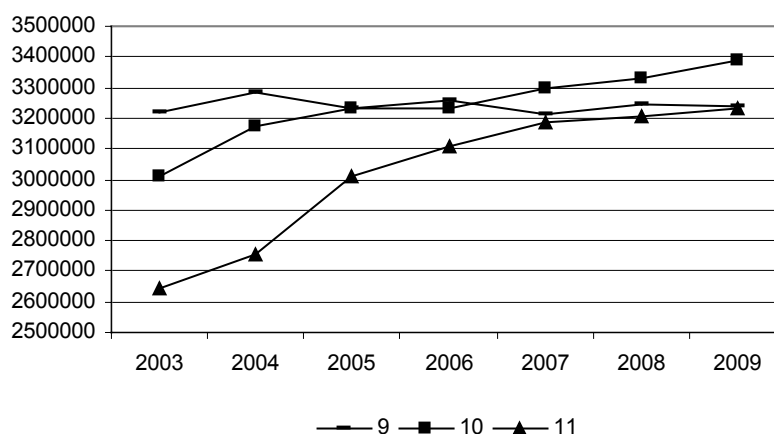
<sup>20</sup> A este respeito ver Mendes (2001).

Gráfico 2: Projeção de matrículas para a 5 a 8ª série do Ensino Fundamental



A trajetória estimada das matrículas para as séries do ensino médio é interessante. Enquanto que a matrícula para o 1º colegial permanece praticamente constante ao longo do período (segundo as estimativas), para as últimas duas séries do ensino médio a tendência é de forte aumento das matrículas, condizente com o aumento da escolarização líquida, já apontado anteriormente. Vale dizer que esse aumento das matrículas no ensino médio pode estar refletindo simplesmente um crescimento do tamanho da coorte escolar ou mudanças nas características dos alunos, que levaram a um

Gráfico 3: Projeção de matrículas para a 1 a 3ª série do Ensino Médio



## Comentários Finais

O objetivo do artigo foi o de apresentar um modelo de fluxo escolar. A partir de dois anos de PNAD (2002 e 2003) identificaram-se as taxas de transição escolar para o ensino fundamental e ensino médio. Com base nessas taxas, expandiu-se a PNAD até 2009, no que se refere às características educacionais. A contribuição do artigo está no fato das taxas de transição serem condicionais ao atraso escolar, diferentemente dos modelos anteriores que supunham um modelo



markoviano. Além disso, como a base de dados é a PNAD, características sócio-demográficas também foram consideradas nas estimativas.

De acordo com as estimativas obtidas, as taxas de ingresso e aprovação diminuem com o atraso escolar e aumentam com a escolaridade dos pais, como esperado. Para a evasão, o resultado é o inverso. Os resultados das projeções de matrículas indicaram o que parece ser um ajuste das matrículas nas séries iniciais ao processo de universalização do acesso a escola que ocorreu na década passada: as matrículas nas primeiras séries estão se normalizando, aproximando-se das medidas de tamanho populacionais das coortes. Para o ensino médio, as projeções indicaram um aumento significativo das matrículas neste nível de ensino, principalmente, para as duas últimas séries do ensino médio, condizente com aumento da escolarização líquida para este nível de ensino. Este aumento tanto pode ter origem em coortes maiores ou em mudanças de características sócio-demográficas.

## Referências Bibliográficas

- BARROS e MENDONÇA. *Consequências da repetência sobre o desempenho educacional*. Série Estudos do Ministério de Educação e do Desporto, Brasília, 1998.
- FLETCHER, P. R. e RIBEIRO, S. C. “*Modeling education system performance with demographic data: an introduction to the PROFLUXO model*”. In: Brazilian Issues on Education, Gender and race. BARRETO, E. e ZIBAS, D. (Orgs). Fundação Carlos Chagas, São Paulo, 1996.
- FILMER, D. e PRICHETT, L. (1998). What Education Production Functions Really Show. *Policy Research Working Paper 1795*, World Bank, Washington D. C.
- FLETCHER, P. H. e CASTRO, C. de M. Mitos, estratégias e prioridades para o ensino de 1º grau. *Estudos em Avaliação Educacional*, v. 8, jul-dez 1993.
- FLETCHER, P. H. e RIBEIRO, S. C.. O ensino de primeiro grau no Brasil de hoje. *Em Aberto*, ano 6, n. 33, jan/mar 1987.
- HANUSHECK, E. A. The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, v. 24, n. 3, p. 1141-1177, Set 1986.
- HANUSHECK, E. A. The impact of differential expenditures on school performance. *Education Researcher*, v. 18, n. 4, p. 45-51, 1989.
- KLEIN, R. *Produção e utilização de indicadores educacionais: metodologia de cálculo de indicadores do fluxo escolar da educação básica*, 1995. (Mimeo).
- KLEIN, R. e RIBEIRO, S. C. O Censo Educacional e o Modelo de Fluxo: O problema da Repetência. *Revista Brasileira de Estatística*, vol. 52. Rio de Janeiro, 1991.
- MENDES, M. Descentralização do Ensino Fundamental: Avaliação de Resultados do FUNDEF. *Planejamento e Políticas Públicas*, n. 24, 2001.
- LEON, F. L. L. de e MENEZES-FILHO, N. A. Reprovação, avanço e evasão escolar. *Pesquisa e Planejamento Econômico*, v.32, n.3, 2002.
- RIBEIRO, S. C. “A pedagogia da repetência”. *Estudos Avançados*, vol. 12, n. 5, 1991.